

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2001, 00:14:33 ; Search time 36.13 Seconds  
(without alignments)  
1439.891 Million cell updates/sec

Title: US-09-494-297-2  
Perfect score: 3945  
Sequence: 1 MKTRFPNKLTNTORVLS.....IAGISLGWGHTRIRKHD 757

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	633.5	16.1	659	S40043	adhesin - Streptococcus
2	603	15.3	638	S54418	fibronectin-binding
3	164	4.2	1039	T30856	protein F2 - Streptococcus
4	159	4.0	1243	S60138	sex factor aggrega
5	157.5	4.0	5005	F82884	hypothetical prote
6	151	3.8	2367	S70172	toxin B - Clostrid
7	149	3.8	1315	T28679	fibronectin-binding
8	148.5	3.8	1883	G82875	hypothetical prote
9	145	3.7	1036	T30311	S-layer protein -
10	144.5	3.7	1092	T30214	fibronectin-binding
11	143	3.6	1185	A42404	collagen adhesin -
12	142.5	3.6	627	T03821	prolipoprotein p65
13	139.5	3.5	888	E82885	hypothetical prote
14	139.5	3.5	1167	F71909	hypothetical prote
15	139	3.5	621	B64546	chaperone and heat
16	139	3.5	1433	A36734	baicillopeptidase F
17	139	3.5	2366	S10317	toxin B - Clostrid
18	138.5	3.5	625	T03837	surface lipoprotei
19	138.5	3.5	2484	T26216	hypothetical prote
20	138.5	3.5	2607	T26215	hypothetical prote
21	137	3.5	1177	I64233	hypothetical prote
22	135	3.4	1117	S33851	fibronectin-binding
23	135	3.4	1176	A33856	surface-layer 125K
24	135	3.4	1228	T40468	surface layer prot
25	135	3.4	1301	S51323	SAC protein - yea
26	135	3.4	1650	T18444	hypothetical prote
27	134	3.4	1668	A60272	Iga-specific metal
28	133	3.4	965	T32574	hypothetical prote
29	132.5	3.4	1116	A35129	surface layer prot

30	132.5	3.4	2287	2	T21312	hypothetical prote
31	132.5	3.4	3097	2	T00021	DN-cadherin - frui
32	132	3.3	1103	2	H82884	multiple banded an
33	131.5	3.3	-867	2	D82940	conserved hypothet
34	131	3.3	843	2	S33442	EF protein - strep
35	131	3.3	1092	2	S42798	fibronectin-binding
36	131	3.3	1822	2	S33441	EF protein - strep
37	130.5	3.3	881	2	S37895	LH1 protein - strep
38	130.5	3.3	908	2	C70168	DNA polymerase I (
39	130.5	3.3	2166	2	G70163	hypothetical prote
40	130	3.3	621	2	A71961	90kda chaperone -
41	130	3.3	1372	2	T43296	cell fusion protei
42	130	3.3	3381	2	T42389	versican precursor
43	130	3.3	5175	2	T20992	hypothetical prote
44	130	3.3	5198	2	T43290	hemisciental precu
45	129.5	3.3	1684	2	S10789	amylase A-180 - al

## ALIGNMENTS

RESULT 1	S40043	adhesin - Streptococcus pyogenes
C:Species: Streptococcus pyogenes		
C:Date: 19-Mar-1997 #sequence-revision 19-Mar-1997 #text-change 28-May-1999		
C:Accession: S40043		
R:Seqs: S.; Avil, A.; Tovl, A.; Bursstein, I.; Caparon, G.; Hanski, E.		
Mol. Microbiol. 10, 1049-1055, 1993		
A:Title: Protein F: an adhesin of Streptococcus pyogenes binds fibronectin via two di		
A:Reference number: S40043; MUID:95020565		
A:Accession: S40043		
A>Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-659 <SEL>		
A:Cross-references: GB:LI0919; NID:9425479; PIDN:AAA26964.1; PID:9425480		
Query Match	16.1%;	Score 633.5; DB 2; Length 659;
Best Local Similarity	26.1%;	Pred. No. 6.1e-30;
Matches 203; Conservative 128; Mismatches 237; Indels 211; Gaps 30;		
QY 20 SKNSKREFTVTVGVFLMIRALVTSWVGAKTVEGVSSSTPMAINPDSSSEYRWYGESEYV 79		
DB 24 TKRRRFAVTVLVGVFPMLLACAGAFGGVAVAAADKTVPSHSP--NPEFPWYGYDAVG 81		
QY 80 RGHPRYKQFRVAHDLRVNLEGRSRYOVCFNLKRAPLGSDSSVKKRWYKHHDSITRFED 139		
DB 82 KEYPGYNITRTHDLRVNLEGRSRYOVCFNLKRAPLGSDSSVKKRWYKHHDSITRFED 141		
QY 140 YANSPRTGDELNOKLRVAVYNGHPQANGIMEGLEPLNARIVTQDAVWYYSNDADISNP 199		
DB 142 YATRTKLGKEELEQRLSLILYNAYPNDANGYMGLEHLNATVQYAVVHYSNDSQYQF 200		
QY 200 DESFKRESNLSVTSQSLSMRQALKQILDNPATKMPQVOPDPLSTFESEDKDKDXN 259		
DB 201 ETLWESAEKGRKISRSQVTLRREALKRLIDPNEAIVANKVIRGLNLFESF-----N 254		
QY 260 KGYONLISGLVPTKPTPGDPPMPNPOTTSVLIRKKAIGDYSKLGGATV-----QLT 315		
DB 255 EATONLISAEVYDPPDKPGE--TSEHNKPTP-----LQGTPLPEPKHP 298		
QY 316 GDNVNSFOARVSSNDIGERIELSDGTYYLTTELNSPAGYSIAETPKYKAGKAVYTIIDG 375		
DB 299 DDNLFTPLPVMID---GEV-----PEVSSSLPALPLMPE-----LDG 337		
QY 376 KQENNKKEIVE-PYSVEAYNDFEESVLTQVYAFYAKNNGSSQVYCCNADLKSP 434		
DB 338 QEV--PEKPSIDPLIEVPRE-----FNNKQOSP 364		
QY 435 PDEEDGKMTPTDFTTGEVY--THIAGRDLEFKYTVKPRDTPDPLFKHKVLEKGYRER 493		
DB 365 LAGE-----SGETFEYITVEYVGNQ-----QNPVIDK-----KLPRNTEGF--S 399		

```

QY 494 GQALYSLSLTFETOLRAANQALAYFTDSALDKKLNKRYH-----FG 536
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 400 GNMVETEDTKEPEVLMGQ-----SESEVFTKQIQTGMSGOTTPOVTEEDTKEPEVLM 453
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 537 DMDNSTLAVAKILVEAYDS-----NPQLTDLDFEIPNNKXOSSLGTOWHPELDVD 589
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 454 GQSES-----VEFTKDTQTGMSGOTTPOIETEDTKEPE-----486
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 590 IIRREDKKEVLPVHNHLBRTVTGAGDRDKDFEFELKNNKO-ELL-----SQTAKTD 644
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 487 -VLMGQSESEVFT-----KDTQTGMSGOTT--QVETEDTKEPEVLMGQSESEVFT 536
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 645 KTNLEFKDGKATILMKHGESLTLQGLPGYSYLVEKTDSEGYKVKVNSQEVANATVSKTG 704
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 537 KDTQTGMSGOTTPOIETEDTKEPEVLMGQSESEVFT-----KDTQTGMSGQSFETA 587
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 705 ITSET-----LAFENNK-----EPVVP--TGVDKINGLYALIVINGIEL 743
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 588 TVEDTRPKLVFHEHNPQKVEENREKPTKNITPLPPTGDIENLAVLGIIILSVSL 646

```

RESULT 2

SS4418

fibronectin-binding protein precursor - Streptococcus pyogenes

C:Species: Streptococcus pyogenes

C:Date: 06-Sep-1996 #sequence\_revision 13-Mar-1997 #text\_change 15-Oct-1999

C:Accession: S54418; A44792; S28448

R:Taylor, S.R.; Valentin-Welgand, P.; Timmls, K.N.; Chhatwal, G.S.

MOL: Microbiol. 13, 531-539, 1994

A:Title: Domain structure and conserved epitopes of Sfb protein, the fibronectin-binding

A:Reference number: S54418; MUID:95089690

A:Accession: S54418

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-638 <TAL>

A:Cross-references: EMBL:X67947; NID:G511149; PIDN:CAA48133.1; PID:G511150

R:Taylor, S.R.; Valentin-Welgand, P.; Jerlstrom, P.G.; Timmls, K.N.; Chhatwal, G.S.

Infect. Immun. 60, 3837-3844, 1992

A:Title: Fibronectin-binding protein of Streptococcus pyogenes: sequence of the binding

A:Reference number: A44792; MUID:92363585

A:Accession: A44792

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 371-638 <TAA2>

A:Cross-references: EMBL:X67947

A:Note: sequence extracted from NCBI backbone (NCBIN:111014, NCBIF:111015)

C:Genetics:

A:Gene: sfb

C:Keywords: fibronectin binding

F:1-49/Domain: signal sequence #status predicted <SIG>

F:50-638/Product: fibronectin-binding protein #status predicted <MAT>

Query Match 15.3%; Score 603; DB 2; Length 638;  
Best Local Similarity 27.4%; Pred. No. 3.7e-28;  
Matches 207; Conservative 118; Mismatches 248; Indels 182; Gaps 32.

```
QY SKSKRFVTLLVGVFLMIFALTYSMVGATYFGLVESSTPNAINDSSSERYMGYESIV 79  
      : :: ||| | ||||| : : : : : : : : : : : : : : : : : : : : :  
Db 17 TKRRKRFVNTLVGVFFMLLASAGAFGCVAAAEKTEVPHHVS - QNEFPFWGYDFX - 73  
  
QY RGHPRYKGFRAAHDRVNLSEGRSYOVYCENLKKAPLGSDSVKKMYKHKGISIKFED 139  
      : :: : : : : : : : : : : : : : : : : : : : : : : : : : :  
Db 74 KG-PTRY --- HNQLNLNNGSKTYQACFNELKRPEPKEGSYFPNNWKAKRMGSSETPFK 128  
  
QY YAMSP -----RIIGDELINOKLARVVMNGHQNMNGINGEPLINAIKVTOEAWYYSD 192  
      : : : : : : : : : : : : : : : : : : : : : : : : : : :  
Db 129 YADNRKNONESRVIDVELEKNIILVLTVNGYPNNNGIMEGLEPLNALIVTQNAAWYYSD 188  
  
QY NAPISNPDSEFRREESUNVTSQSLSMRQALKOLIDPLATKMKROYDDPOLSTFESE 252  
      : : : : : : : : : : : : : : : : : : : : : : : : : : :  
Db 189 NSSIINTNFETFEKEDLNIRKEPDSLMNVAAKTLDIPDLTSEESLKPVASTIRLFEEQS 248
```

Qy	253	DKGKMYKGVONLLSGGLVPTKPTPGDMPNPOCOTSVLIRKXALGDSKLLGATL	312
Db	249	D-----KLYONLLSAFVFNPKPGE--TEPHGKPTL-----DGPILBGP--	290
Qy	313	QLTGDVNSFOARVFSNDIGERIELSDGTLYLTELNSPAGYSIAEPITFEVAGKYTTI	372
Db	291	QRPNSELEPTLPV-----MIDG-OEVEPEVSESLPALPMLPELDGGEVEV	338
Qy	373	-----IDGKOI-ENPMKE--IYEPYSVEAYNDFEESVLTQONAKFYA	414
Db	339	PSESLPALPMLPELDGGEVPEKPSVDLPEDR-----	373
Qy	415	KNKNGSSOVVYCFNADLSKSPDSEDEGKTMTBDPTTGVEKY--THIAGRDLEKYTKPRDT	473
Db	374	-----YEFNNKQOSPLAGE-----SGTEYITVEYNGO-----QNPVDI	407
Qy	474	DPDTFLKHKKVYIEKGYREGQALIEYSGLETQRLAATYLTQDASLEDKOKLYNH	533
Db	408	DK-----KLPNPTGF--SGNNVETEDTRKEPVLMSGO-----SESVETPKDQOTGMS	452
Qy	534	GFGDMNDSTLAVAKILVEYAKDSNRP-----QJLDLDFEIPNNKNYOSLIGTOWHPEDL	587
Db	453	G-----QTTQPVGE-----TEDTRKEPVLMSGGSESVETPKDQOTGMSQOTASVETEDT	501
Qy	568	VDI-IRMEDKEVLPVTHNLTLKRYTGLAGDRTKQFHFHEILKNNKOKELLSQYTKTBT	646
Db	502	KEPGLVMSGGSESVETG-----KDTQOTGMSGQOTTP-----QVEETDTRKEPVLMSGGSE--	550
Qy	647	NLEF-KDGKATINLKHEESTLLOGEGSYLVAKETDSEGYVAKYNSOANATVSKGTGI	705
Db	551	SVEFTPKDQOTGMS--GFSETVITVEDTRKPLVHFPHDNNPEKVEERKPT-----	598
Qy	706	TSDETLAEFNKKEPVVP--TGVDOKINGYLTALYIYA	739
Db	599	-----KNITPLPATDQIEVLAFLGJILLIS	624

RESULT 3  
T30856 Protein F2 - Streptococcus pyogenes  
C:Species: Streptococcus pyogenes  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000  
C:Accession: T30856  
R:Jaffe, J.; Natanson-Yaron, S.; Caparon, M.G.; Hanski, E.  
Mol. Microbiol. 21, 373-384, 1996  
A:Title: Protein F2, a novel fibronectin-binding protein from Streptococcus pyogenes  
A:Reference number: Z20907; MUID:97011581  
A:Accession: T30856  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1039 <JMF>  
A:Cross-references: EMBL:U031980; NID:g1654115; PID:g1654116; PIDN:AMC44522.1  
C:Genetics:  
A>Note: PRUF2

```

Query Match          4.2%; Score 164; DB 2; Length 1039;
Best Local Similarity 19.7%; Pred. No. 0.087;
Matches 171; Conservative 124; Mismatches 279; Indels 296; Gaps 46;

QY  57  STPNAINP--DSSSEYRWYGESYVGHPPYKQFVAHDLRLNLEGSRSYQYVCFNLKRAF 115
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 110 NSPLAIKGRKDATHQLTLYFTNYING--LDRVQJSAELSLPLENKEYLE----- 156

QY 116 PLGSSSVAKWKYKKHGGISTKFEEDYAMSFRTIGDELNOK--LRAYMNGHPONANGIMEG 173
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 157 -----NTNISDEKST--TGGGEITRYKGTAVNLVYGESTSTESNYITNG 196

QY 174 L-----EPLNAIRVTOGAVMYVSDGNAPIS-----PDSEFRRESSENLVSTS 215
      -|-----|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 197 LSNVGGSLSYNT--ETGEFMYVYVYNPRKRIPIPAVLNLMGCFARTRKQGGENDSNLSVSA 254

```

[illegible]

A:Gene: clua  
 F:1-43/Domain: signal sequence #status predicted <SIG>  
 F:44-1243/Product: sex factor aggregation protein #status predicted <MAT>

Query Match 4.0%: Score 159; DB 2; Length 1243;  
 Best Local Similarity 19.6%; Pred. No. 0.23;  
 Matches 160; Conservative 100; Mismatches 243; Indels 314; Gaps 42;

QY	98	LEGSRSYQYCFNLKRAFLPGSDSSVKWKYKKKKDQISTKPEFYAMSFRITGDELINOKLRA	157
DB	335	MAGSLNYS-----KKTIDGGILPTISDGNQHSVS--FDN-----TL	370
QY	158	VMYNGHPQANANGIMEGLEPLNLAIRVTOEALVWYISDNAPISNPDESEKRESESNLVSTSL	217
DB	371	VVGSPYPPSSGVHVKVS-GPTSGANATGDS--WSSNPPTNETWASATAYIDYKAIGPS--	424
QY	218	SLMRQALKOLIDPNLNTK-----MPKQVPPDPOLSTIFESDEKDKYKQGNL--LSGL	270
DB	425	-----LDVTEMDVGTGKMSWGAMNLPKDGSTISVTMGTTD-----ANMALLINGQ	471
QY	271	VPTKPTPPDPMPMPQOPQTTSVLRKRYAIGDYSKILEGATLTLDGDNVNSFGARVSSN	330
DB	472	LENKITTTP--EPRIIPKPRATYYDDATPTDNTK-----AVTQTDGDLN-----	516
QY	331	DIGERIELSD-GTYTLTLEINSPAGYSIAE-----PITFKVEAGKVYTI-----ID	374
DB	517	--GALVNKQETERNWVLSNEVLPAGHEVIKSYVMTDPLPEGFKILDEQSKTSLSPDYDLTFD	574
QY	375	GK-----QIENPKKEIPEPYSEKAVNDPEESVLTTONYAKFYAKKNGSSQV	423
DB	575	EKTNTVTLTAAYKATLEAMKDLNQAAYQVPRK-----TLQGVV-----TKDGS--	617
QY	424	VYCFENADLKSPDSEDDGKTMTPDFT--TGEEVYTHIAGRDLEFKYTVKPRPTDPTDLKH	481
DB	618	---FKKDL-----ETLINDITYNSNEVE-----VHTDPRPREKSNEN	651
QY	482	IKKVIKEGYREKQGAIE-----YSGLTEQ--LRAATOLAIYYFTDSA-	522
DB	652	-----ASGTTINGQIGDVNATNYYKLLMDLSGKGIASSKEDIVRG-----FYVDAAP	700
QY	523	---ELDKDKLKDYG-----FGDMNDSTLAVAKIYE-----	551
DB	701	DVAVDVLKNIYSKDSQGEKYGITAKVYSSVKAPAEVQVFLADAKIAPKGFVYVSVD	760
QY	552	-----YADSSNPQLT-----DLDFEIPNNKYYO-----SLI-----	578
DB	761	PQTFYTNVYQGTGNNVITDPTMFKEGASGAYQNTDQIDEGNSIEGDTVKNNTVPRKVK	820
QY	579	-----GTOMH--PEDLVADIIRMEDKKEVIFVTINLTLRKTVTG-----	614
DB	821	QYSVDSGKTKWHDSKDLRPTDSDSNYDKLDFEFTANGDYTKILLDGNFESSQWDLAKAKVT	880
QY	615	-----LAG-----DRTKDFHEIEELKNNKQELLISQYVTKDTN-----	647
DB	881	DRKDNDAIQCFKVLNASKGRDVTKDFNNHVFQKDEKKEVLIQIIFTPKDISITSLASNDP	940
QY	648	-----LEFKGKATINLKHGESLFLQGL--PEGSYLVAKFEDSDEGKYKVNVSQEVANA	698
DB	941	DRLLITLTMSFKD-----VTLKGGTGAELANYL--DKESKIVAPNIGOL--D	982
QY	699	TVSKTGITSDET--LAFENNKEPVPTGYDOKINGLY	733
DB	983	TTSRT--VTGDNTKDKDKITKSNTVKVIRPOLRPMINKTV	1018

RESULT 5  
 F82884  
 hypothetical protein UU495 [imported] - Ureaplasma urealyticum  
 C:Species: Ureaplasma urealyticum  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C:Accession: F82884  
 R:Glass, J.T.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit

A:Reference number: A82870

A:Accession: F82884

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-5005 <G1A>

A:Cross-references: GB:AE002148; GB:AE222894; NID:g6899495; PIDN:AAE30907.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: U0495

A:Genetic code: SGC3

Query Match 4.0%; Score 157.5; DB 2; Length 5005;

Best Local Similarity 19.8%; Pred. No. 2.5;

Matches 158; Conservative 100; Mismatches 285; Indels 253; Gaps. 36;

5 REPNNKLTNTQORVLSKNSKREF-----VTL-----VGVEIMIFALVT 42

1598 RIKGQNIISIPATISDDDKRYTKCMATNLGLDYEIVSVIKKQTKNVSPIIFELPS 1657

43 SNVGAKT--VEGLVESPPNAINDSSEIRWYGESIVRGHPYTKQPRVAHDLRVNLE 99

1658 GPICGYTLAPYFKITSLW-----DPYHHEKKNADUKINLR 1694

100 GSRSGYOVCFNLKKAFFLGSDSSVKKWK-KHDGISTKPEDYAMSPRTGDELNOKLRAV 158

1695 -IENIGASLFN-----KDIKIFRKNKGKQISITHKVT- 1728

159 MYNGHPONANGIMEGLEPLNAIRVTOEAVWYSDNAPISNPDESEFKRESESNIVSTSOLS 218

1729 -----ANANVEMERKDLLRNREYTLERVY-----LKNKPDNGSVSES----- 1768

219 LMRQALKQILDRNLATKMPKQYDPDFOLSFESSEKGDKNMGYOLLSGGLVPTPKPT 278

1769 -----DYDLMWIEMLNKTFR-----LLPTKPLGI 1793

279 GPPMPNPQPTTSVLRKRYAIGDYSKL-EGATLQLTGDNVNSFOARVSSNDIGE-- 334

1794 IGAPRIEISDNGAKVOL-KFALNDFPDVAKENOTPKF--NIOPNNANLNEISEHSG 1849

335 RIELSDG-TYTTLELNSPAGYSIAEPIFEKVEAGKYTTIDGKQIENPKKEIVE-DYSVE 392

1850 KYEIIDGKFEFVANLNN-----IKVKNKEYKY--KIYEDENQDVGNGYKIN 1894

393 AYNDEE-----FSVLTQNTA--KEYYAKNKGSSQVYCFNADLSPDSDEGK 442

1895 FRNDYKEPNNVYDASVNTQTYVFNKFAISFSNNLTDVANKONISINDS---R 1950

443 TMRPPTTGEVYKTHAGDLFEYTVKPRDTPDFTLKHKKVIEKGYREKGOAIEYSG 502

1951 VETIQGYHFRKATISDNKRYWNTI-PAPTN-----IGNKKNIALNFE-L 1995

503 TETQL---RAATQALAIYFTDSALDKLKDYGFGDMNDSTLAVAKILVEYAQDSNP 559

1996 NQNGILSNLTYFAALY--SKDVNVDE-----NHANVYIKNVNPDQISTPK 2042

560 QLTDLDFPIRNNKYOSLSGTQNHPRDLDVITMEDKKEYIY-----THNLTLRK 610

2043 SSIVYDLKQMDENKITISLLHSNDQI--FEDKNHMLKIAKISIDELAHDOITNS 2098

611 TYTGLAGDRTKDHFETELKNNKQELISQTVKTDNTLFEKDGKATINLK----- 660

2099 T-----THDYDLLEKENNEMLKLTQILN-KERTKYRVKQVFTSKRPSDTIYNGV 2149

661 -----HGESLTLOGLEPGYSYLVKETDESGYKVKVNSQEVANAV--SKTGITSD 708

2150 KQIIVYRINNHNNTIDKLEKSTLNSVKPT-----KTNKKNESIKLWGFNKTGSLE 2204

709 ETILA---FENKPEV 720

2205 NKYAKLVYKDNNOOI 2220

# RESULT 6

70172 Clostridium difficile

toxin B - Clostridium difficile

C:Species: Clostridium difficile

C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 15-Oct-1999

C:Accession: S70172; S44271

R: von Eichel-Streiber, C.; Meyer zu Heringdorf, D.; Habermann, E.; Sartling, S.

Mol. Microbiol. 17, 313-321, 1995

A:Title: Cloning in on the toxic domain through analysis of a variant Clostridium dif

A:Reference number: S70172; MUID:96079281

A:Accession: S70172

A:Status: Preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-2367 <ON>

A:Cross-references: EMBL:Z23277; NID:g761713; PIDN:CAA80815.1; PID:g761714

A:Experimental source: isolate 1470

R: Sartling, S.; von Eichel-Streiber, C.

submitted to the EMBL Data Library, July 1993

A:Description: Cloning and sequencing of an Clostridium difficile toxin B mutant.

A:Reference number: S44271

A:Molecule type: DNA

A:Residues: 1-1323, 'N', 1325-2367 <SAR>

A:Cross-references: EMBL:Z23277

C:Superfamily: cpl repeat homology

C:Keywords: cytotoxin

Query Match 3.8%; Score 151; DB 2; Length 2367;

Best Local Similarity 19.9%; Pred. No. 1.9;

Matches 171; Conservative 94; Mismatches 292; Indels 304; Gaps 42;

5 REPNNKLTNTQORVLSKNSKREF-----SKNSKRFYTLVGVFLMIFALVTSNMGAKTVGLVE--SS 57

1568 KPMNRKGSNTSDSLMSFLSESNIKSIFVNFLOSNIKFIIDANFLISGTTSGGEFICD 1627

58 TPNAINP-----DSSSEYRWY-GYESYVRGHPYKQFRAVHADLVNLEGSSTOYCFN 110

1628 ENNNIQPYFIKFNLTETNTLYVGNRQNMIVEPN-----DIDDSGDISSTVIN 1676

111 LKAPFLGSDSSVKKWKYKHHDSITKFEYAMSPRTGDELNOKLRAVYNHPO-----N 166

1677 FSQKLYLYGIDSVNK-----VVISPNLYDEIITVYEETNNTPPEYIYD 1722

167 ANGIMEGLEPLNAIRVTOEAVWYSDNAPISNPDESEFKRESESNIVSTSOLMQLKQ 226

1723 ANIYNEKIN-VAINDLSTIRYV-----SNDGNF-----ILMSTSEKNKVSQVKIR 1767

227 LI---DPMIATKMPKQYDPDFOLSFESSEKGD-----KYKGYQNLISG-- 268

1768 FVNVFVKDKTILANKLS-----FNFSKQDVPVSEIILSTPSY---YEDGLIGD 1813

269 -GLVPTKPTPGDPRMPNPQPTTSVLRKRYAIGDYSKLLEGATLQLTGDNVNSFOARV 327

1814 LGLV-----SLYNEKFTYNNGMWVSG--LIYINDSLYFRPPY- 1850

328 SSNDIGERIELSDGYTTLELNSPAGYSIAEPI-----TKVEAGKY 370

1851 -NNLITGFYTVGDKKYFNPINGAA-SIGETIIDDKNYFNQSGVLQGVSTEDGFKY 1908

371 ---TI---IDGKQIENPKKEIVEP---YSVEAYNDFEFSVLTQNTQ-----AKFY 413

1909 FAPANTIDENLEGEALIDFGKLIIDENIYFEDNYRGAAVEWELDEMYFSPETGAKR 1968

414 AKKNGSSQVYVCFNAD--LKSPDSEDDGKTMPTFTGGEVYKTHIAR----- 461

1969 GLMQIDDK--YFNSDGVQKGFVINDKHYFDDSGYKQVGYTIDGKHFFAENGEM 2026

462 -----DLFKYTVKPRDTPDFTLKHKKVIEKGYREKGOAIEYSGLETQLRAATOL 513

2027 QIGVFNTEDEGFKY-----FAHHNEDL-----GNEGEIEISYGLINFRNK----- 2066

[illegible]

```

RESULT      7
T28679
fibrinogen-binding protein homolog - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C:Accession: T28679
R:Joseffson, E.; McCrea, K.; Nl Eldhan, D.; Cox, J.; Hook, M.; Foster, T
Microbiology 144, 3387-3395, 1998
A:Title: Three new members of the serine-aspartate repeat protein multigene family of S
A:Reference number: Z20510; M01D:99096700
A:Accession: T28679
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1315 <OS>
C:Cross-references: EMBL:AJ005646; NID:e1318791; PID:e1318792; PIDN:CAA06651.1
C:Genetics:
;Gene: sdrD

```

```

Query March 3.8% Score 149; DB 2; Length 1315;
Best Local Similarity 20.2% Pred. No. 0.98;
Matches 203; Conservative 93; Mismatches 336; Indels 372; Gaps 45;

QY      8 NKLTLNTQRLVLSKNS--KRETVTLGVGVLFMLFVLVTSMWGAKTVFGLVSESTP----- 59
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      208 NSNNENNADILLPKSTAPKLNTR-----MRIAIVQPSSTEAKNVNDLITSNTLTLVDA 262

QY      60 ---NAINP--DSSEFRRYGYESTYVRGHPYTKQRVANDLKVNLKESNSTOYCEPNLKA 114
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      263 DKNKKIVPAQOYLSLKSQITVDVKVSGDYF-----TIKYSPTVOYVGLNPEDI 311

QY      115 FPLGS---DSVKKWYKKHDG---ISTKPEDYAMSPRITGDELNKLRAVMTNGH--- 163
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      312 KNIGIDKPNNGENIATKHDFTANLLITFTDY-----DRENSQMKINISYIDA 364

QY      164 ---PONANGIMEGLEPLNAIRVQEAWVYYSDDNAPISNPDESFKRESESNL----- 211
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      365 DTIPVSKNDVEFNTVIGTTTKTTANIOYPDYVVENKKSISAFETVYSHVGNENPNQY 424

QY      212 -----VSTSOELSRQALK-----QLIDPNLATKMKPVDPDFQSLISESDKGGKYNKY 262
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      425 KOTIVNSENSESLNAKLVQAVYHSSYPNNIGQINKDDTD--IKITV-VKRGITLNGY 480

QY      263 QNLSGLVLPTRPPTGGDPMPMPNOPOTTSVLIRKRYAIGDYSK----- 305
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      481 D-----VNTKELT-----DVTNÖYQIKITFYGDNNNSAVIDEGNADSAYVMVN 522

QY      306 -----LLEGATLQLTGD-----NV 319
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB      523 TKFOYTNSESPTLVQOMALVSTGKNSVSTGNALGFTNNOSGAGQOEYVYKIGNYWEIDTK 582

```

```

OY 320 NFO-----ARVSSN---DIGRIELSDOTYLLTEL-----NSPAGS 355
Db 583 NCVQELGECGVGNVTYVTDNNNTKVGAEVTKEDSTYLPNLPNGDYRVERSNLKEJTE 642
OY 356 IAE-----358
Db 643 VHPKSGONNEELDSNGLSSVITVNGKDNLSADLGICYKPRYNLGDYVWEDTNKNGIDODE 702
OY 359 -----PITFKYBAGKY---TIIDG---QIENPNKEI-----VEPPSVEAYNDE 397
Db 703 KGISGVTVTLKDENGNVLTQVTTDADGKRYFTDLDSNGNRYKEFTPEEGVPTPTVTSGSDI 762
OY 398 EEFIS-VLTQNT-----AKERYAKNNKSSQVYVCFMNDLSPDSED---CGKT 443
Db 763 EKDSNGLTGVYINGADNMTLDSGFYKTPKRYNIGNVWEDTNKDGKO--DSTKGISGVAT 820
OY 444 MPDPTTGEVKKYTHIAGRD-LFKYTVKPRDTPDPTFLKHKKVIEGYREKGOAIEY---499
Db 821 YLLKNGNEGVLOTTKTKDKRGKYOF-----GLENGYKVEFEP 859
OY 500 SGLTETOLRAATOLAI-YFTDSEAEIDKDK-----LKDY-----HFGS 536
Db 860 SGTPTQVQSGSTDEGIDSGNSTTGVIKDKDNDTIDSGFYKPRYNLGDYVWEDTNKNGYO 919
OY 537 DMNDSTLANA-----KIIVYEQDSNPR-QLDLDBFEPNNKXQSLQTOHNEP 586
Db 920 DKDEGISSVTVTLDENDKRVLTQVTTDENGKYOFPDL-----NNQSTYK-----963
OY 587 LVMDITIMEBKKEVIP--VYHNLETKR-----TWTGL---AGDRTKDFHEFIELKNNKOEL 636
Db 964 -----VEFEPISGTPRISVSGNDEKDSNGLTITGVYIKADNMTLDSGF---YKTPKYSL 1016
OY 637 -----LSQIVTKDTKTNLEFKDGNKATINLKHGESLTLQGLPBGSYLVYKENDSEGYK 669
Db 1017 GDYVWYDSNKKDKOSTEKIGIKDVKVTLLNEKEGEVIGITKTEDENGKYCFDNLDSGKYKVI 1076
OY 690 VASQEVAAATVSKTGITSDETLAEFNKKEPVYPTGVQDKI-NKY 732
Db 1077 F--EKPRAGLTQGTINTEDKADGGEVVTYITDHDFTIDNCEY 1118

```

```

RESULT      8
GB2875
hypothetical protein U0556 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: G62875
R:Glass, J.L.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.,
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum; Alternate views of a
A:Reference number: A82870
A:Accession: G62875
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1883 <GUA>
A:Cross-references: GB:AEO02154; GB:AF222894; NID:g6899557; PIDN:AAF30971.1; GSPDB:G
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: U0556
A:Genetic code: SGC3

Query Match          3.8%; Score 148.5; DB 2; Length 1883;
Best Local Similarity 19.1%; Pred. No. 1.9;
Matches 155; Conservative 102; Mismatches 276; Indels 277; Gaps 34;

QY   103 SYOVCFNLRKRAF--PLGSDSSVKKWKKHDDIGSTGFEDYAMS-----143
    ::||| |:::||:: ||::| |::| |
Db   951 NNDTSFENLKTFTFDHPHMFITSNEKRKAERITDAMFETPPDENISKLDEKANASSVPY 1010

QY   144 ---PRTGDDELNR-----LRVMYNGHGPNANGIMEGLEPINAIRVTOEAVWYS 191
    :|::| |::| |::| |::| |::| |
Db   1011 KKVLDSOTGEDINTNKTFDFDKFYALLTLIQYNASKRTMWNKSSEEVSKLDPELI--EIKKY- 1067

```



A:Title: A Fibrinogen-binding protein of *Staphylococcus epidermidis*.

A:Reference number: Z20781; MUID:98261511

A:Accession: T30214

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1092 <NII>

A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 3.7%; Score 144.5; DB 2; Length 1092;  
Best Local Similarity 18.0%; Pred. No. 1.4;  
Matches 159; Conservative 145; Mismatches 269; Indels 309; Gaps 44;

```
QY 8 NKTINTLNTQR-VLSKNSK-----RFTVTLGVLFMLFALVTSVNGAKTYVGL-----V 54
DB 3 NKNKNLITTKKPKLPKSNKYNKPKFTVGTASI-----VIGATLLEGLGHNEAKAE 52
QY 55 ESSTPNAINDSSSEKRWYGESEYVGHPPYKQFVAHDLRLVLEGSRSYQVYCFNLKKA 114
DB 53 ENSVQDVKDSNTDLS-----DSNDQSDDEKNDVYINNOSINTDNOQI-----IK-- 100
QY 115 FPLGSDSVKWKYKHKGIDGISTKFEDYAMS-----PRITGDELNOKLRVAVYNG 162
DB 101 -----KEETNNYDGLKEKRESDRETESTYNDENAEFLAQKTPQDNTHLLEEVYESS 151
QY 163 HPQANGIMEGLEPLNAINTQEAFAVYVYSGDAPISN-PDESEKRESESNVYSTSLMR 221
DB 152 SVSSSSSIDTAAQPSHTTINRESEVQTSNDVEDSHVSDPANSKIKESNESKEENTTE 211
QY 222 QALKQIDPLNATKMPKQVDDFOLSFESSEDKDKRYNKGYNLSSGLVPTKPTPGDP 281
DB 212 Q-----PNKV-----KEDSTTSQPSGYTNI-----DE 233
QY 282 PMPNPQPTTSVILIRYVAGDYSKLEGLATLQITGDNVNSFQARVSSNDIGERIELSDG 341
DB 234 KI-SNDELMLNPINEYE-NKARPLSTTSAPSIRKRYVNOALAEQGSVNMHLIKVTQD 290
QY 342 TYVLTFLNSPAGSIAE-----PIF-----VEAGKYVTI-IDGKQI----- 378
DB 291 SIT-----EGTDSEGVYKADAEANLIYDVEVDKKSQDYMIVTDKKTVPISDLT 343
QY 379 -----ENPKIEVEPYS-VEAYNDEFEFSVLITQNYAKKYYAK 415
DB 344 DSFTTIKIDNSELITATGIDYDKNKOITFTFDYDKYENIAHLKITS-----YIDK 397
QY 416 NK-NGSSQVYVCENADLSPDSEDKGKTMPDF-----TTGEVK-YTHIAGDLFK 465
DB 398 SKVPNNNTKLDVETKALSSV-----NKTIYEQRPNEPNTANLQSMPTNI--DTKN 448
QY 466 YVYKPPDTPDPTFLKIKKIEKYGREKQALFYSGLTE---TQLRAATQALYYFTDSA 522
DB 449 HNY-----EQTIIYNPLR-----YSAKETNNVNSGDEGSTIIDSTIIKYYKKGDDN 497
QY 523 EL-DKDKLKDYGFGDMNSTLAVAKILVEYAODSNPQOLDLDFIPNNKKYQSLIGTQ 581
DB 498 NLPDSKRITDISEYEVTD-----DYAO-----LGNNDVYNINFGNI 535
QY 582 WHEDELVDIIRMED-KKEVIPTVNLILKTYTGLAGD----- 618
DB 536 DSRPY-IIVKYSKDYDPKDDYTIQQTVMQTITNEYGERFASYNONTIAFSTSSGGGQ 594
QY 619 -----RTKDPHE-----TELKNNKOELL-----SQYKTKTKLNEK 651
DB 595 DLPEKTYTKIGDVWEDVDKDGIONTNDNEKPLSNVLVLTYPDGTSKSVYTD-----E 648
QY 652 DGRKATIN-LKHG--ESLFTLOGLPEGSYLVKET-----DSEGYKRVNSQEVANATV-- 700
DB 649 DGRYQDGLKNGLTLYKITE-TPEGYTPFLKHSGTNPALDSBGSNVAWYTIINGODDWTIDS 707
QY 701 -----SKTGITTSDE-----TLAFEN 715
DB 708 GFYQTPKYSLSGNVWYDTNKKDGIQGDDEKGISGVKVTLDKEN 749
```

RESULT 11

A42404 collagen adhesin - *Staphylococcus aureus*

C:Species: *Staphylococcus aureus*

C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 10-Nov-1995

C:Accession: A42404; S27665

R:Patil, J.M.; Jonsson, H.; Guss, B.; Switalski, L.M.; Wlberg, K.; Lindberg, M.; Hook

J. Biol. Chem. 267, 4766-4772, 1992

A:Title: Molecular characterization and expression of a gene encoding a *Staphylococcus*

A:Reference number: A42404; MUID:92165839

A:Contents: FDI 574

A:Accession: A42404

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1185 <PAT>

A:Cross-references: EMBL:M81736

A:Note: sequence extracted from NCBI backbone (NCBIP:83982)

Query Match 3.6%; Score 143; DB 2; Length 1185;  
Best Local Similarity 19.7%; Pred. No. 1.9;  
Matches 151; Conservative 108; Mismatches 279; Indels 230; Gaps 40;

```
QY 99 EGSRSYQVCFNLK-----KAPLGSDDSVKWKYKHKGIDGISTKFEDYAMSPLITGDE 150
DB 269 QGTGYNSSINIKYTKITNEOKEFVNNSOA-----WYQEHG-----KEEYNGAS 313
QY 151 LNOKLRAVYNGHPQANGIMEG-----LEPLNAIRVTOEAVYVYSGDAPISNPDESFK- 204
DB 314 FNHTV-----HNINANGIEGTAVKGLVKLQKQMDTK-----APIAN-VKFKL 355
QY 205 RESESNVYSTSLMRQALKQIDPLNATKMPKQVDDFOLSFESF-----DKGDKY- 258
DB 356 SKKDGSVYKDNQ-----KEIEIITDANGIANIKALPSGDYILKEIAPRPTEDKKEYP 410
QY 259 -----NGYONLSSGL-----VPTKPTPGDPMPNPQPTTSVILIRYVAGDYS 304
DB 411 FYMKDITDNOGFTFTIENAKAIKTKDVSQAQWEGTQKAP-----TIFKLIKQDDNO 464
QY 305 -----KLEGATLQITGDNVNSFQARVSSND-IGERIELSDGYTLTELN----- 349
DB 465 NTTPVDAKELKLEDDGTTKTWSNL-----PENDKNGKALK-----YLVEVANAQGE 511
QY 350 --SPAGSIAE-----PI-TFKEAGKYVTIIDGKQIENPKIEVEPYSVAYND 396
DB 512 DTPPEGYTKKENGLVYVTEKPIETTSISEKVDKDKNDODGKRPER-----VSANLLAN 566
QY 397 FEEFSVLITQNYAKFY-----AKNNSSQVYVCENADLSPDSEDKGKTMPPTTG 451
DB 567 GERVKTLDVITSETNMKYEERKDLPRYDEG-KKIEYVTEHVKDYTTDINGTITNNKYTPG 625
QY 452 EVKYTHIAGHDLFKRYVYKPPDTPDPTFLKIKKIEK-----GYREKG 494
DB 626 ETSATYTKMNDNNNDGKRPTIEIKVELYQDGKATGKTALINSNMWTHWTGLDEKAG 665
QY 495 QALEYSGLTETQIARAQ-----LAIYFTDSAELODKLKYDHGFGDMNSTL 543
DB 686 QQVKYVVEELTKVKGTYTHVNDNMGNLIVTNKYTPETTSISEKGYWD-----DKDN-- 737
QY 544 AVAKILVEYAODSNPQOLDLDFIPNNKKYQSL-----IGQWHPEDLVYIIRMEDKKEYI 600
DB 738 -----QDGKREKSVN-LIADGEKVTLDVTSFTNMKYE-FKDLPKYDEKKTI- 784
QY 601 PYVHNLTFRKTYVGLAGDPTKDPHEFE-----LKNKKOELLSQLYKT-DKTNLE----- 649
DB 785 -----EYIV--EDHVKQDITDINGTITNNKTPPGSTATYVKKMNDNNNOGKRPE 833
QY 650 -----FKDGKAT-----INKHGESLTLQGLPE-----GYSYLVE-TDSEGYKRVKN 691
DB 834 TEIKVELYQDGKATGKTALINSNMWTHWTGLDEKAGQOQVYVVEELTKVGYTTHVD 893
QY 692 SQEVANATYSK-----TGITSDETLAFENKKEPVPYTPGVQDKI--NG 731
```

Db 894 NNDMGNLIVTNKYPETTSISGEKVDKDNODGKRPEKYSVNLNG 941

RESULT  
T03821

prolipoprotein p65 - Mycoplasma hyopneumoniae  
 N:Alternate names: major immunogenic surface Lipoprotein  
 C:Species: Mycoplasma hyopneumoniae  
 C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 07-Dec-1998  
 C:Accession: F03821  
 R:Haradiri, M.B.; Kim, M.F.; McIntosh, M.A.; Wise, K.W.  
 submitted to the EMBL Data Library, July 1997

Query Match	3.68;	Score 142.5;	DB 2;	Length 627;
Best Local Similarity	20.58;	Pred. No. 0.74;		
Matches 163;	Conservative 116;	Mismatches 265;	Indels 251;	Gaps 41;

QY	19	LSKSKRPTVTVGVFLMIRPALVTSWGAKTVEGLVSSSTPMANIDPSSSEIRMYGYEST	78
Db	1	MKKARKP-LBLTSLTLLAPFVFTLLTSAGC---LOKNSLSEVNTALADSLTAGFNEE	56
QY	79	VRGHPYKOFVADLRYNIEGSRSOYV-CFNLKAAPGLGSDSVKWKWKHDGIS---	134
Db	57	T---YRDPGTLDKOGLNSG-QSYAFYFYLYOK---LKNLSLV---SYDNLASIG	102
QY	135	-----TFEEDYAMS--PRIT---GDELNOQLRAVMYNGHPQANGIMEGLEPL	177
Db	103	TTEENMWLYLNPRTKYPNGKMSDNLPLVNYSGNEKYNEIGSVFGDFNKDSYGLAEVYKKA	162
QY	178	NAIVTQEAVMYVSODNAPISNPDSFRESBSNLVSTLSQLSIMQALKOLIDPLALAKMP	237
Db	163	NLTMWSYGA-----NDPELAIFNEFKMA-----SIIKPSSEAKKLIDPA---ERA	206
QY	238	KQVDDPOLSLIFESEDDGDKXNKYQNL-----LSGGLVPTAPPTPGDPMPBPQ	287
Db	207	NFLAEKGLMLKAEVNNKRIEINTLNDMLIKELALNPKLSINLVGYLPLPMSG-----	258
QY	288	PQTTSVLIRKAIADYSKLEBGLATLQLTGDNVNSFQARVSSN-----DIGERIELSDGT	342
Db	259	---FIKLKYLTLTYAK-IETDPEINELPEKINKIIRETAIKKNVNYIDVYDKSIMWSD	313
QY	343	YTLTELN-----SPAGY-SIAPETPFVEAGKYVTLIDGOIENPKEIYEPVSVEAND	396
Db	314	KNLAKNNDPFIHSIQGIKKKIHOLLKL-----TLDOEKSDSNAE--ELKNTTFED	364
QY	397	FEFESVLTQNYAKFYAAKKNNGSSOYVYCFNADLKSPPSEDDGKTMTDPDFTTGKYKT	456
Db	365	FDE-----NKPTYSKYV-----	376
QY	457	HIAORLEKYVRKRDNDPDLFLKHKKYVLEKGYRENGOALIEYSGLLETOURAATOLAIY	516
Db	377	-----DLSVFAKSNKEPELEK-LNEKKQTSSEFLAQKST-----	407
QY	517	YFTTSAEELDKRLKDYHGFGDMNSTLAVAKILVEYADSNSPOLTLDLDF--FIPNNKY	574
Db	408	FDTQGEAIAKD--DKRTFGNI-----VREIV-----SLPFDNFDPRELLIPVKNPF	451
QY	575	OSLIGTQWHPED---LVDIIRMEDKK-----EVIPTVHNLTTRKVTYGLAGDRTKD	622
Db	452	VKAIINSYLGKPAAGSLIKIDIEOLENKVADARPNIKIFDTIILDSFIKRWMAFFA-----E	506
QY	623	FHEIETELKNKNO--ELLSQYKVT-----DKTNLEFWDG---KATINLKGESLITDGLP	671

```

Db      507  LNTDQEIKEFMSQIIFLTLRNAILSPFDTKV--KDSATFKILMLKPEQIITLLGLG 564
Oy      672  EGYST-----LVKEDSEGYKVKVNSOEYANNTVSKTGISDSETLAFENKPEPV 721
Db      565  KTPSPVKPEKPKDDGSMQIUTSSOK-----QET-----STGST-DSTATITENQPAE 612
Oy      722  PT-----GVDQKIN 730
Db      613  QTNSEDSSTSDSKSN 627

```

## RESULT 13

E82885  
 hypothetical protein YU481 [imported] - Ureaplasma urealyticum  
 C.Species: Ureaplasma urealyticum  
 C.Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C.Accession: E82885  
 R.Glass, J.L.; Lefkowitz, E.J.; Glass, J.S.; Helmer, C.R.; Chen, E.Y.; Cassell, G.H.  
 Submitted to GenBank, February 2000  
 A.Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a  
 A.Reference number: A82870

Query Match	3.58	Score 139.5	DB 2	Length 888
Best Local Similarity	18.48	Pred. No. 1.9		
Matches 150	Conservative 125	Mismatches 256	Indels 283	Gaps 39

[illegible]



```

Db      505 I EONOKIILFKGIDPIESTETTKTNDLRRR-----I I K I K I D---I K D 546
Qy      538 MNDSTLAVAKIIVEAADSNPQ-----L I-----D L D F P I P N--N N K Y--O S L I G T O W 582
Db      547 Y I N G D C-----R L V L E Y R N N N S T S K E V S L T L N K N O E D S V N I E M U L T Y N R E V T L N I K I N H 602
Qy      583 H P B D L--V D I I R N E D K E V I--P Y T H N I L T R K Y T T G L A G R T K D F H E I E L K N N K O E L L S Q 639
Db      603 N T N N I S I D I S K I N D K T F V E P S V T R K L T N F K N I T N F-----D E R K O I W S R 647
Qy      640 T V K T D T N L E F K D G K A T I N L K H G E S I L T L O G L P E G S Y L V K E T D S E G Y K V K V N S O E V A N A T 699
Db      648 I M E P T I N D D E M F S O S L N-----S T K I R A I V E N L S--T N K I D N P I I K I--V N D N G K P K L Q 700
Qy      700 V S K T G I T S D E T L A F E N N K E P V P T G V D O K I N G Y L 733
Db      701 I H I N I N I N D-----Q N N E Y K R I V K F E N K N S Y L 730

```

## RESULT 14

```

F71909
Hypothetical protein jhp0613 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: F71909
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Meberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; M0ID:99120557
A:Accession: F71909
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1167 <ARN>
A:Cross-references: GB:AE001493; GB:AE001439; NID:g4155161; PIDN:AAD06194.1; PID:g415516
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0613

```

```

Query Match      3.5%; Score 139.5; DB 2; Length 1167;
Best Local Similarity 19.1%; Pred. No. 3;
Matches 136; Conservative 110; Mismatches 282; Indels 185; Gaps 32;
Qy      103 S Y O V Y C F N L K A P L G S D S S V K W-----Y K K H D-----G I S T F E D Y A M S P R 145
Db      504 S Y I A L I N I T Q A--A Q N R D S L K N F K N I A L T D S I D L E K K D K G V I F G F E L F E D L K E N K E 562
Qy      146 I T G D E L N Q L R A V Y N G H P Q N A N G I M--E G L E P L N A I R V T Q E A V W Y Y S D N A P I S N D E S 202
Db      563 I K T T M E Q N I R V I I--G N P Y S S G A K S E N D N Q N L S H P K L E K R Y E T Y G K N S T A Q N----- 616
Qy      203 F K R E S E N L V S T Q S L S M R A L K-----Q I D P M L A T K M P Q V D D Q L S F E S 251
Db      617--K S T R D T L I H S--I R R A S O L L A D K G V L G F V N G S F I D S K S A D G F R C V A Q D F S--H L Y A L 672
Qy      252 E D G D Y K N K G Y O M L S G L V P T K P T P G D P M P N Q P O T S V L-----I R K Y A 299
Db      673 N L G N M A R T S E E R-----K K O G C I P S G S R N V T A I I F P V K D A P N H T I P Y E 721
Qy      300 I G D Y S K L E G A T I O L T G D N V N S F O A R V F S S N D I G E R I--E L S D G T Y T L T E L N S P A G S I A E 358
Db      722 V E D L K R E A K L N I L A N F E N D S V P F K E I T P N D G D W I N G N D E F E K L I P L R D K S K I F N 781
Qy      359 P I F E K V E A G V Y T I I D G K O L E N K E I V E P Y S V E A Y N D E E F S V L T Q N A K A Y Y A N K N 418
Db      782 A I--F D L N S N G V K T S R D-----P W-----V Y N F S Q K T L M O S V O N----- 813
Qy      419 G S S O V V Y C F N A D K-----S P D S E D G K T M P D T T G E V K Y T H I A G R D L F E Y T V 468
Db      814-----C I D T Y A D A L K R F E R N E R F A F K O R T A K D K I K S A D-----R K H L N D R E I T T D K T 863

```

```

Qy      469 K P R D T P D T L K H I K K Y I E G Y R E K G A I E Y S G I T E T O L---R A A T O L A I Y-----F T D S 521
Db      864 K I A M T D G-----L K N--K L I-----K E N L P E S G M E R V R A L Y R F N K O W L Y M K N L I N R Q 912
Qy      522 A E L D K--D K-----L K O Y H G F G M N D S T L A V A K I L V E A A D S N P O L T D D F I P N N K Y 574
Db      913 S O L P K I F P D K S A R N V Y I N T V G N K D F S A L V D F I S D F I S--P N Q A Y P L Y I D D L G N R Y 971
Qy      575 O S L I G T O W H P E D L V D I I R M E D K K E V I P-----V T H N I L T R K Y T T G L A G R T K D F 623
Db      972 N A I S G Y-----A L N L F R R Y K D N A I T E E I F Y Y I A I F H K G Y L E K Y K N S L A K E A P R-- 1023
Qy      624 H F E I E L K N N K O E L L S O T V K T D K T N L F E K D G K A T I N L K H G--E S I T L O G L P E G S Y L A Y K E T 661
Db      1024---I A L S E D F E K E L S M U G K E L A E L H L A N E S G E M H T S V K H N L E S A E M E G Y---Y D V I O M K R 1077
Qy      682 D S E G Y K V K-----V N S O E V A N A T Y S K T G I T S D E T L A F E N N 716
Db      1078 D K R G D R I K Y H H I T T O I P K K A F D Y V N G K S A I D W I E R Y O I T R K D S L I E N N 1130

```

## RESULT 15

```

B64546
Chaperone and heat shock protein C62.5 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 20-Aug-1999
C:Accession: B64546
R:Tombs, J.F.; White, O.; Keriavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; M0ID:97394467
A:Accession: B64546
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-621 <TOM>
A:Cross-references: GB:AE000541; GB:AE000511; NID:g2313299; PIDN:AAD07278.1; PID:g231
C:Superfamily: heat shock protein 90

```

```

Query Match      3.5%; Score 139; DB 2; Length 621;
Best Local Similarity 21.9%; Pred. No. 1.2;
Matches 130; Conservative 85; Mismatches 227; Indels 152; Gaps 30;
Qy      201 E S F K R E S E N-----L V S T S O L S M R O A L K O L D P N L A T K M P K O V P D O F O L S F E S E R K G- 255
Db      28 E I F L R E L V S N A S D A L D K L N T L M L T D E R L K-----G L N T T P S I H L S P S O K K T Y L T I K D N G I 82
Qy      256--D K Y N-----K G Y O N L S G G L V P T K P P T P G D P M P N Q P O T S V L I R K Y A I G D Y 303
Db      83 G M D K N D I E H L G T I A N S G T I N P L S A-----L S G D-----K K D S A L I O G F O G Y G F Y 127
Qy      304 S K L E G A T I O L T G D N V N S F O A R V F S S N-----D I G E I E I L S D G T Y T L T E L N S P A G Y---- 354
Db      128 S A F M V A S K I Y V Q T R K V S D A Y A V M S D G K G F E I S E C V K D E O G E T I L F L K D E D S H P A S R 187
Qy      355-----S T A E P T F V E A G K Y T T I I D G--Q I E N P K E I V E P Y S V A Y N D F E F----- 400
Db      188 W E I D S V V K Y S E H I P F P I-----F L T Y T D T H E G D G O K K E I K E B K C Q I L Q A S A L M K M K 243
Qy      401 S V L T T O N Y A K F Y A K N K N G S S O V V Y C F N A D L S P D S E D G K T M P D T T G E V K Y T H I-- 458
Db      244 S E L K D K Y K E F Y O S F A D N E P L S Y I N K-----V E G S L E Y T L F Y 284
Qy      459-----A G R D L K-----Y T V K P R D T D P-----T F L K H I K V I E K G-----Y R E K G 494
Db      285 I P S T A P D M E R V Y K S G V K L Y V K R V F T T D D K E L P S Y L F E V G V I D S E D L P L N V S R E-- 342
Qy      495 Q A I E Y S G I T E T O L R A A T O L A I Y F T D S A E L D K I R K O Y H G F G M N D S T L A V A K I L V E--Y 552

```

```

Db 343 --LQOKKILANISASVKKI--LSEIERLSKDB-KVYHNFYE-----PFGVKMBGLX 392
QY 553 AQDSNPQOLDLDFIPNNKQYSLGTQWHEBDLVDIRMEDKEVIYPT-HNLIRKLT 611
Db 392 GDFENKKEITLLETFYSKDK--EKLISLKEYNEL-----KENQSIYYLIGENLILKA 444
QY 612 VTGLAGDKTFHFEIELKNNKQ--LLSQVTKDKNTLEFGKQKATINLKHGSLTLOG 669
Db 445 SPLL--EKYAAKGVDVLLSLSEIDAFWPGVNEVDKT--PFKQA-----SHSESLKEIG 494
QY 670 LPE-----GYSYLKETDSEGGKRVVNSQEVANATVSRGTGISDETLAENN 716
Db 495 LEEIHDEKQDFKDLMAKFE--ENLKDITKGVLLSHLTSVALIGDQNNMMAN 547

```

Search completed: June 7, 2001, 00:18:18  
Job time: 225 sec